

Generalization of the PageRank Model

Noah E. Friedkin¹

January 21, 2014

¹ *Center for Control, Dynamical Systems and Computation, College of Engineering, and Department of Sociology, College of Letters and Science, University of California, Santa Barbara .*

Abstract

This paper develops a generalization of the PageRank model of page centralities in the global webgraph of hyperlinks. The webgraph of adjacencies is generalized to a valued directed graph, and the scalar dampening coefficient for walks through the graph is relaxed to allow for heterogeneous values. A visitation count approach may be employed to apply the more general model, based on the number of visits to a page and the page's proportionate allocations of these visits to other nodes of the webgraph.

1 Introduction

In a multi-agent network, the total influences of the agents are related to the number and length of the walks in the network. A model of this relationship [1] is

$$\begin{aligned}\mathbf{V} &= (\mathbf{I} + \alpha\mathbf{W} + \alpha^2\mathbf{W}^2 + \alpha^3\mathbf{W}^3 + \dots)(1 - \alpha) \\ &= (\mathbf{I} - \alpha\mathbf{W})^{-1}(1 - \alpha)\end{aligned}\tag{1}$$

where $\mathbf{W}_{n \times n}$ is row-stochastic and $0 < \alpha < 1$ is scalar. The v_{ij} of $\mathbf{V} = [v_{ij}]$ correspond to the relative net influence of agent j on agent i . Thus,

$$\mathbf{r} = \frac{1}{n}\mathbf{V}^T\mathbf{1}, \quad \mathbf{r}^T\mathbf{1} = 1,\tag{2}$$

is the average relative net influence of agent j on all agents of the system, and may taken as a measure of the centrality of node j in the system [2, pp.1485 -1487]. The average may be based on the n values of each column, or $n - 1$ values when the main diagonal values of \mathbf{V} are excluded. If the model is expressed as follows [3]

$$\mathbf{r} = \frac{1 - \alpha}{n} + \alpha\mathbf{W}^T\mathbf{r}\tag{3}$$

then equation 2 is its solution.

2 Generalization

The generalization [4] of equation 2 is

$$\begin{aligned}\mathbf{V} &= (\mathbf{I} + \mathbf{A}\mathbf{W} + (\mathbf{A}\mathbf{W})^2 + (\mathbf{A}\mathbf{W})^3 + \dots)(\mathbf{I} - \mathbf{A}) \\ &= (\mathbf{I} - \mathbf{A}\mathbf{W})^{-1}(\mathbf{I} - \mathbf{A})\end{aligned}\tag{4}$$

where \mathbf{A} is diagonal with each value $0 < a_{ii} < 1$. Clearly, this model includes the special case $\mathbf{A} = \alpha\mathbf{I}$, and its corresponding centrality values are also given by equation 2.

An agent with a_{ii} near 0 or w_{ii} near 1 is in a receiver (sink) state with outflows near zero. An agent with a_{ii} near 1 and w_{ii} near 0 is in a transmitter (flow-through) state. A useful implementation of the model couples a_{ii} and w_{ii} . The simplest coupling is $a_{ii} = 1 - w_{ii}$.

With this coupling assumption, an implementation of the model may be based on the proportion of visits to a page i that terminate on it, w_{ii} , and the proportions of visits to a page i that activate a hyperlink to page $j \neq i$, w_{ij} . If such a specified \mathbf{W} is in hand, then \mathbf{A} is empirically determined and its values reflect the extent to which a page is a sink or transmitter. Thus, the model moves from an adjacency matrix basis of centrality toward a basis that attends to the behaviors of users within the structure of the webgraph.

3 Discussion

This paper develops a generalization of the PageRank approach. The assumption of a scalar dampening coefficient is relaxed to allow for heterogeneous values. A visitation count approach may be employed to apply the more general model. The construct \mathbf{W} , which is conventionally based on the binary hyperlink structure of the webgraph, need not be so restricted. Further note that the necessary information to implement this approach is node-local, i.e., based on the number of visits to a node and the node's proportionate allocations of these visits to other nodes. A node's allocation weights w_{ij} , ($j = 1, 2, \dots, n$) may be selectively monitored for stability and updated.

References

- [1] Noah E Friedkin and Eugene C Johnsen. Social influence and opinions. Journal of Mathematical Sociology, 15:193–206, 1990.
- [2] Noah E Friedkin. Theoretical foundations for centrality measures. American Journal of Sociology, 96:1478–1504, 1991.
- [3] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems, 30:107–117, 1998.
- [4] Noah E Friedkin. A structural theory of social influence. Cambridge University Press, New York, 1998.